Reply to reviewers and editors:

We thank all reviewers for their careful reading of the manuscript, and for their many constructive feedbacks. The original comments by reviewers are in black font, our replies are in blue.

Reviewer #1

This paper presents the formulation and sensitivity analysis for the inverse modeling of SO2 and NOx emission over China using satellite data. While the authors seem to emphasize that the joint assimilation saves 50% of the computational time than assimilating SO2 and NO2 separately, the benefit of the joint assimilation should be more than that. This needs to be clarified.

Thanks for your comment. Yes, joint inversion should have more benefits in addition to save computational time. We have added the discussion below to Sect. 4.3.

When evaluating with OMI retrievals, joint inversion shows better results than separate inversion for SO_2 or NO_2, but not both, depending on the value of γ . When γ is 20, 50, or 100, NO_2 NCRSME for E-joint-dγ improvement appears to be smaller than that for E-NO_2, but SO_2 NCRSME for E-joint-dγ is larger than that for E-SO_2. Conversely, when γ is 1000, 1500, or 2000, SO_2 NCRSME for E-joint-dγ is smaller than that for E-SO_2, but NO_2 NCRSME for E-joint-dγ is larger than that for E-NO_2. This is similar to the findings by Qu et al., (2019b) in which the months when joint inversion show better result than separate inversion for SO_2 (NO_2) have worse result for NO_2 (SO_2). The benefit of joint inversion for improving only one species is similar to Qu et al. (2019b) and is likely to due to the complicated relationship between these two species through different chemical pathways. For example, O_3 and OH are key species that connect the chemistry of SO_2 and NO_2 and aerosols can affect the photolysis and heterogenous chemistry. Hence, while joint inversion to improve both species can not be demonstrated here, it should be reviewed as the first step of simultaneously assimilating multiple species (including AOD, NH_3, and other trace gases) to optimize emissions. Until then, the system is not ready to holistically evaluate the benefits of joint assimilation to improve the model in a systematic manner. It is worthy noting that Xu et al. (2013) showed the feasibility of using MODIS cloudfree radiance to optimize emissions of SO_2 and NO_2 at the same time. Future research should add the aerosol optical depth or visible reflectance (as well as tropospheric O_3 if reliable) as constraints to further evaluate the benefits of joint assimilation for improving model overall performance in a systematic matter.

In the paper, γ is introduced to balance the SO2 and NO2 terms. In theory, it is not needed if the uncertainty terms can be well quantified. The optimal value of γ is determined pretty arbitrarily. There are objective ways (such as Hollingworth-Lönnberg and NMC methods) to determine the observational errors and its covariance terms instead of relying on arbitrary balancing.

Thank you for the comments. In a practical way, we still need use γ to balance the SO_2 and NO_2 terms no matter how observational errors are quantified, as "When it is not balanced, SO_2 observations have very little impact on the inversion results as the optimization algorithm will firstly minimize the observational term for NO_2 unless many more iterations than is computationally feasible are performed, which is caused by the fact that observational error and valid number of NO_2 observation are respectively smaller and larger than the counterparts of SO_2**.** We thus **subjectively** derive γ in a **non-arbitrary** way in order to focus equally on both species, thereby tackling the imbalance in their observational constraints. In this manner, the cost function is defined to server the purpose of joint inversion of SO_2 and NO_2 emissions" "Similar balance approach that adjusts contribution of observation terms in the cost function is used in the past work that assimilates satellite trace gas retrievals to invert emissions (Qu et al., 2019b) or invert the aerosol optical properties from skylight polarization measurements of AERONET (Xu et al., 2015).." We have added the quoted text to Sect. 3.2.1.

Actually, OMPS SO_2 and NO_2 observational errors are quantified in an **objective** way. In a clean region (such as equatorial Pacific ocean (10°S–10°N,120°W–150°W) that is far from emission sources), true SO_2 (NO_2) concentrations should be zero or negligible, while both negative and positive retrieval values exist. Thus, it is reasonable to use the variance of SO_2 (NO_2) retrievals over the clean region to represents SO_2 (NO_2) observation error variance. This error estimation approach is widely used in trace gas retrieval research community (Li et al.,

2013;Yang et al., 2013). These precision values can be used as the observation error in the cost function of data assimilation. However, we should notice that the estimated observation (retrieval) errors only represent the observation error distribution of the products a whole; it cannot represent the observation error distribution for every pixel, because the pixel-level error is amenable to spatiotemporal change of cloud fraction, satellite observation geometry, aerosol impacts, etc. In theory, if the uncertainties can be analytically described at the pixel level, they would be directly applied to improve the satellite product in the first place.
We have add the discussion above in Sect. 2.1.

We also have added the text below in Sect. 3.1
In the optimization formulation, the forward model errors are also considered as part of the observation error term. However, while several ways to construct model error covariance matrix exist, including the Hollingworth-Lönnberg (Hollingsworth & Lönnberg, 1986) and NMC (Bannister, 2008) methods, their application for off-line CTM model error characterization deserves a separate a study. The Hollingsworth method extracts observation error variance (including forward model error) from (observation – background) covariance statistics with the assumptions that observation error is spatially uncorrelated, background error is spatially correlated as a function of distance, and observation error and background error are uncorrelated. The assumption that background error is spatially correlated as a function of distance only is suitable for the meteorological fields that vary smoothly, but for chemical species, emissions also contribute significantly to model errors and emissions are spatially correlated. The NMC method is normally applied to weather forecast models or on-line-coupled weather-chemistry models (Benedetti and Fisher, 2007). Off-line CTMs such as GEOS-Chem use the meteorological reanalysis and so, NMC is not applicable here to quantify the CTM's transport error. Consequently, CTM's transport errors are neglected in the past emission optimization work (Wang et al., 2016) and are adopted in this study. Admittedly, this simplification should be studied in future together with the evaluation and developments of methods to characterize off-line CTM errors.

Specific comments:

Line 215: It is surprising for the authors to choose less than 3% reduction in the cost function between two iterations as a criterion to halt the minimization. The L-BFGS-B can be very slow. Such a condition can often terminate the minimization prematurely. This needs to be changed if it is not a typo.

This is a good point. L-BFGS-B can be very slow to converge. But at some point we run into practical limitations on the amount of time we can spend iterating, thus we choose less than 3% reduction in the cost function between two successful iterations as a criterion to halt the minimization. Based on the criterion, for E-SO_2 and E-NO_2, we picked 5th and 6th iteration result, respectively. Further tests show that the more iterations (after <3% reduction of cost function) doesn't yield discernable difference in the cost function values (Fig. S3) and optimization results (Table S1 and S2). We have added the figure and table below to the supplement and added corresponding description at last paragraph of Sect. 3.1.

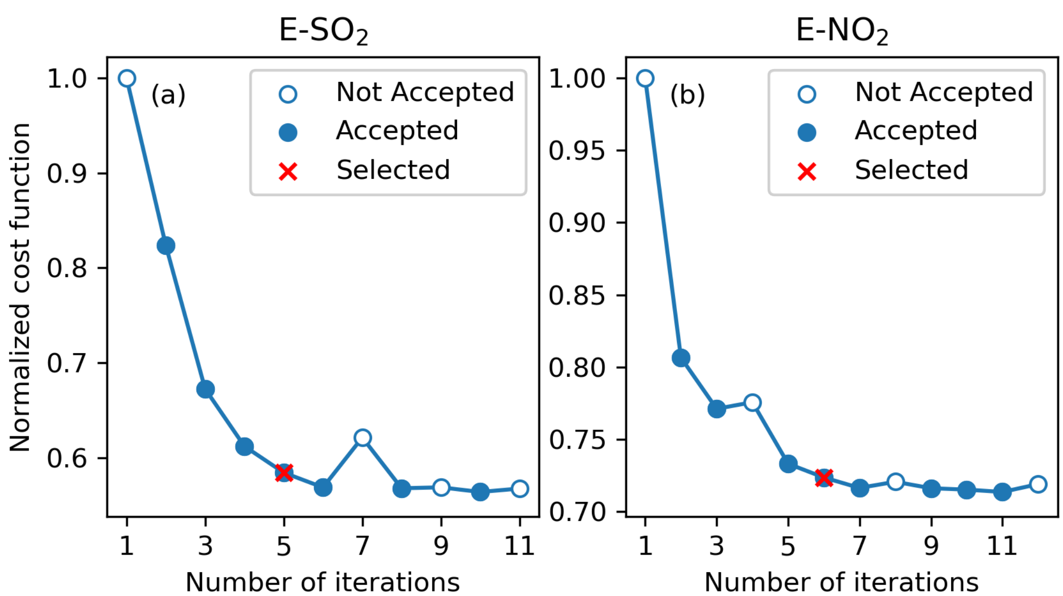

Figure S3. Normalized cost function (the ratio of the cost function at a iteration to that at the 1st iteration) for E-SO_2 (a) and E-NO_2 (b). A iteration is accepted (solid circle) if its cost function value is smaller than that of any previous iterations, otherwise not accepted (empty circle). The 1st iteration (prior) is defined as not accepted. The iterations that are selected based on the halt criterion are marked with red cross.

Table S1. Anthropogenic SO_2 emissions for October 2013 from E-SO_2 at each iteration.

Iteration	1 (prior)	5 (selected)	6	7	8	9	10	11
SO_2 [Gg S]	1166	748	744	Not accepted[a]	746	Not accepted	749	Not accepted

[a]Posterior emission total amount at the iteration that is not accepted (cost function value is not smaller than that of any previous iterations) is not shown.

Table S2. Anthropogenic NO_x emissions for October 2013 from E-NO_2 at each iteration.

Iteration	1 (prior)	6 (selected)	7	8	9	10	11	12
NO_x [Gg N]	714	672	667	Not accepted[a]	666	666	666	Not accepted

[a]Posterior emission total amount at the iteration that is not accepted (cost function value is not smaller than that of any previous iterations) is not shown.

Lines 234-7: Is it really beneficial to balance the cost function this way? Can the SO2 observation errors be objectively determined?

As we address that comment about using γ to balance observation term, SO_2 observation error is **objectively** determined, but the spatial balance problem still exists. Similar balance approach that adjusts contribution of observation terms in the cost function is used in the past work that assimilates satellite trace gas retrievals to invert emissions (Qu et al., 2019b) or invert the aerosol optical properties from skylight polarization measurements of AErosol RObotic NETwork (AERONET) (Xu et al., 2015). Thus we think the subjective but non-arbitrary balance approach should be acceptable, although it is a break from strict Bayesian derivation of the cost function. We have added how the SO_2 observation error (0.2 DU) are determined in an objective way in Sect. 2.1 and the justification of the balance approach in Sect. 3.2.1.

Line 247: It is not accurate to say "emissions are adjusted mainly at locations where prior emissions are large". If there are non-zero emissions, the adjustments can be made. The limitation of using scaling factors is that zero-emission grid points cannot be modified.

Yes. We agree that "The limitation of using scaling factors is that zero-emission grid points cannot be modified." We have added statement of the limitation in the manuscript.

Technical correction:

Line 161, “Terrain reflectivity less than 30o”: Which angle does "Terrain reflectivity" refer to?

This is a typo. We have change it to “terrain reflectivity less than 0.3”.

Line 171: GOES-FP -> GEOS-FP

Corrected.

Line 433: compared the latter -> compared with the latter

Corrected.